Identifying unknown soft coral species of *Anthelia* (Octocorallia) with multilocus DNA barcoding

Mannraj Singh Pataria

Dr. Joshua Miller and Dr. Ross Shaw

Biological Sciences Department, MacEwan University

Contact information: patariam2@mymacewan.ca

Abstract: There are an estimated 1-9 million species of corals yet to be discovered. *Anthelia* is a species of soft coral that belongs to the *Xeniidae* family within Octocorallia. The *Xeniidae* family of soft corals are of interest due to their ability to rapidly recolonize disturbed reefs, which have become more prevalent with global warming. Octocorallia also contains some of the most valuable corals used in jewelry. Identifying corals not only contributes to its conservation and our knowledge of its evolution, but also prevents fraudulent coral jewelry and the overharvesting of coral beds. However, morphologically identifying corals is very difficult and is further exacerbated with global warming or when it is polished and carved into jewelry. Instead, multilocus DNA barcoding can utilize the genetic material of corals to reveal an accurate classification of species and prevent its exploitation. Specifically, genetic loci in the mitochondrial or nuclear genes can be used to tag and classify corals, with referencing done to genetic databases such as GenBank or NCBI. We identified the soft corals *Anthelia glauca, Sarcophtyon trocheliophorum,* and a *Sinularia spp.* that was mistaken for a species of *Nepthea.*

Key words: DNA barcoding, Octocorallia, coral reef, *Anthelia*, mtDNA

Introduction

Coral reefs are often referred to as the rainforests of the sea (Knowlton, 2001). The significance of coral reefs resides more in the plethora of species that reside with or within the corals, rather than the corals themselves. We have barely scratched the surface for identifying corals and there are an estimated 1-9 million species residing in these reefs. In general, there are hard and soft corals found throughout these reefs. Our interest is in soft corals, which lack the rigid calcium carbonate skeleton found in the hard corals and do not form reefs themselves. Soft coral species are spread throughout the world and are more specifically found in the Indian ocean, which is the second-largest coral reef province in the world (Jahajeeah et al., 2021). Soft corals are also a

diverse source of various marine products that provide antimicrobial, anticancer, and anti-inflammatory functions (Jahajeeah et al., 2021). Around 60% of these products also have medicinal potential. In terms of classification, our genus of interest (*Anthelia*) resides in the anthozoan subclass octocorallia and includes some of the most visible and structurally significant sessile macroorganisms found on shallow, tropical reefs and in the deep sea (McFadden et al., 2019). These soft corals are quite common in their shallow-water reef communities and are of interest due to their ability to rapidly recolonize disturbed reefs (McFadden et al., 2019). It is this capacity of recolonization that can be problematic, potentially making them invasive. Identifying these invasive species is important in monitoring coral reef health and diversity. This importance has caused research to shift from morphological to molecular identification; researchers use the coral's mitochondrial genes and nuclear DNA sequences to provide a clear and consistent resolution on its identification (McFadden et al., 2019).

In 2003, researchers proposed the idea of DNA sequences being used as barcodes (Neigel et al., 2007). DNA sequences are to be used to identify species, akin to the concept of having barcodes on retail products. The central concept of DNA barcoding is the utilization of a standard sequence that can correspond to a single homologous gene region (gene region that can indicate a common ancestral origin). This method of using a single gene region is referred to as single-locus barcoding. This "taxon barcode" was developed to circumvent the issues presented by the morphological identification of soft corals and the decreasing availability of taxonomists. In the field, morphologically identifying all the hybrid, cryptic, and inconspicuous species in reefs is an arduous and tedious process that impedes our understanding and classification of corals. This morphological problem resulted in a region of 658 base pairs in the mitochondrial cytochrome c oxidase 1 (COI) encoding gene being proposed as a standard barcode for identification in animals

(Valentini et al., 2009). An ideal DNA barcoding system must have a barcoding marker that is variable, standardized, phylogenetically informative, extremely robust, and short in length. In other words, the gene region should be concise and consistent within a species but varied between species. The target region should also be standardized and contain highly conserved priming sites and reliable DNA amplification/sequencing. However, an ideal DNA barcoding marker that follows all these requirements is yet to be found and the COI encoding gene that was proposed as a standard barcode has come to reveal several issues. The mtDNA (such as the COI gene) can have varying rates of mutation that can result in insufficient discriminatory power (Stemmer et al., 2013). In the case of anthozoan cnidarians, the unusually slow rates of mitochondrial gene evolution have caused the species-specific DNA barcodes (such as the ones in *Xeniidae)* to fall behind other groups (McFadden et al., 2014). This slow rate can leave little room to differentiate between species as they can still share similar mtDNA sequences. Another prominent issue with this single-locus barcoding is that it does not always work with species that are separated by short divergence time or that contain introgressed (transferred) genes from closely related species (Liu et al., 2017). For example, *Anthelia glauca* and *Anthelia rosea* are found in the temperate region of Japan and are sister groups (very closely related) (Koido et al., 2019). Around Indonesia, we find *Anthelia phillippinense*, *Anthelia ternatana*, and another unidentified *Anthelia* species. These species can be separated by short divergence times and are difficult to identify morphologically or with single-locus DNA barcoding, as they can still share similar sequences of DNA if they recently diverged. Lastly, the mitochondrial markers only reflect the maternal lineage of the soft coral, which is often in disagreement with a species' history (Quanttrini et al., 2019)

To circumvent the issues brought forth by single locus DNA barcoding, recent studies have identified multilocus barcodes that distinguish morphospecies of octocorals with a 70–80%

success rate (McFadden et al., 2014). The base concept of multilocus DNA barcoding is using a combination of DNA sequences from varying independent genetic loci to accurately identify species. The main idea is that two species sharing alleles across multiple different genes becomes increasingly unlikely. Research has shown that species become more distinct from one another when multiple, independent loci are used in identifying them. More genetic loci will result in greater separation, precision, and discriminatory power within or between species. Specifically, a barcode that combines cytochrome oxidase I (COI), an adjacent intergenic region (igr1), octocoral-specific mitochondrial mutS gene (mtMutS), and a fragment of the nuclear 28S ribosomal RNA gene has been shown to distinguish MOTUs (molecular operational taxonomic units – an index for classification) that agree with >70% of morphospecies identifications in octocorals. For example, two individuals could have the same COI sequence and still be different species if we analyze other, more differentiable DNA sequences (Liu et al., 2017).

To quantify our multilocus DNA barcoding results, we must analyze our molecular data. Researchers utilize a percent divergence threshold, which is the percent divergence of a gene being used as a barcode (Quattrini et al., 2019). In other words, to what degree do the sequences of our unknown soft coral deviate from the sequences that are known? Databases such as NBCI or GenBank provide these known sequences and generally include DNA sequences of identified species (Neigel et al., 2009). We compare our sequences to this database and look for discrepancies and quantify how much our sequences differ (percent threshold divergence). The common percent divergence threshold for octocoral species detection in barcoding studies is varies and is still a topic of debate (McFadden et al., 2011). Once we have determined the species, we must classify it. Classic biodiversity indices, such as the Shannon index, Simpson's index, are generally used to analyse a species (Neigel et al., 2007). Since these indices rely upon traditional

methods, such as morphology, we need a specialized index in our case to utilize in DNA barcoding. We can use these morphological indices together with molecular operational taxonomic units (MOTUs), which is the biodiversity index for the barcoding method (McFadden et al., 2019). MOTUs utilize the relative abundance of each type of DNA sequence instead of the relative abundance of each species. This allows us to better map out the distribution of a species, such as soft corals, using the molecular data. With this in mind, we will be designing a multilocus DNA barcode to identify and classify unknown species of soft corals - including *Anthelia* - by comparing it to candidate sequences (classified species of soft corals) found in genetic databases.

Materials and Methods

DNA Extraction

We utilized a Zymo animal and a plant DNA extraction kit (Quick-DNA Universal and Plant kit) to extract DNA from *Anthelia spp*., *Sarcophyton spp., Ricordea spp., and Nepthea spp.* These two kits were used to see if one provided a better quality of DNA over the other, given the varying proteins and cells in these corals that needed to be lysed and dissolved. These coral specimens were located on the MacEwan University campus. For DNA extraction, we cut pieces of each coral specimen and measured them to be ≤ 25mg and added 95 uL of water and 95uL of solid tissue buffer, with 10uL of prepped proteinase K. These samples were then incubated at 55C for 3 hours. We then added 2 volumes of genomic binding buffer to these samples (400uL to 200uL) and transferred the samples into separate zymo spin columns and collection tubes. These samples were centrifuged at ≥12000 x g for 1 minute, after which the contents in the tube were discarded. 400 uL of DNA prewash buffer was added to the column and placed into a new tube and centrifuged at ≥12000 x g for 1 minute, with the elutant emptied. 700 uL of g-DNA wash buffer was then added and centrifuged at ≥12000 x g for 1 minute, with the elutant emptied.

Finally, we added 200uL of g-DNA wash buffer and centrifuged at ≥12000x g for one minute and transferred the column to a new microcentrifuge tube and eluted with 50uL of DNA elution buffer, which was then incubated at 55C for 1 min. For the Quick-DNA plant kit, we followed the provided protocol and added up to ≤ 25mg of our coral samples to a ZR BashingBead™ Lysis Tube. We then proceeded to add 750 µl of BashingBead™ Buffer to the tube and centrifuged at ≥10,000 x g for 1 minute. Afterwards, we transferred 400 µl of supernatant to a Zymo-Spin™ III-F Filter in a collection tube and centrifuged at 8,000 x g for 1 minute. We then subsequently discarded the Zymo-Spin™ III-F Filter and added 1,200 µl of Genomic Lysis Buffer to the filtrate in the previous collection tube and mixed. We then transferred 800 µl of this mixture to a Zymo-Spin™ IICR column in a collection tube and centrifuged at 10,000 x g for 1 minute. The flow-through was discarded from the collection tube and the previous step repeated. We then added 200 µl DNA Pre-Wash Buffer to the Zymo-Spin™ IICR column in a new collection tube and centrifuged at 10,000 x g for 1 minute. After this step, we added 500 µl g-DNA wash buffer to the Zymo-Spin™ IICR Column and centrifuged at 10,000x g for 1 minute. The Zymo-Spin™ IICR columns were then transferred to a clean 1.5 ml microcentrifuge tube and 100 µl of DNA Elution Buffer was added directly to the column matrix and centrifuged at 10,000 x g for 30 seconds to elute the DNA. The Zymo-Spin™ III-HRC filter was then placed in a clean collection tube and 600 µl of prep solution was added and centrifuged at 8,000x g for 3 minutes. This eluted DNA was transferred to a prepared Zymo-Spin™ III-HRC spin filter, which was placed in a clean 1.5 ml microcentrifuge tube and centrifuged at exactly 16,000x g for 3 minutes.

Gel electrophoresis

Gel electrophoresis was performed on each DNA extraction and PCR to confirm results. We used a 2% gel with 50mL of TBE buffer and 1 gram of agarose. This was combined in a flask and microwaved until the agarose had dissolved and the solution was clear. 2uL of SybrSafe was added once the solution had cooled to room temperature and then placed in a gel mold. The gels were run at 130v until the samples were about 2/3 down the gel.

Measuring DNA concentrations

DNA concentrations were measured on a Implen NanoPhotometer N60/N50 by using 2uL of water as a blank and 2uL of sample to measure. Data from the Nanophotometer was recorded and added to an excel document.

PCR

We amplified the extracted coral DNA with a PCR designed from Mcfadden et al. (2011) and Koido et al. (2022) in a BioRad T100 Thermal Cycler. The primers COII8068F (5'CCA TAA CAG GAC TAG CAG CAT C3'), COIOCTR (5'ATC ATA GCA TAG ACC ATA CC 3'), ND42599F (5'GCC ATT ATG GTT AAC TAT TAC 3'), MUT3458R (5'TSG AGC AAA AGC CAC TCC 3'), 28S-Far (5'-CAC GAG ACC GAT AGC GAA CAA GTA-3'), and 28S-Rar (5'-TCA TTT CGA CCC TAA GAC CTC-3') were used for amplification. These primers were used at a 10uM concentration and were paired, representing the COI, ND4MUT, and 28s loci respectively. A master mix was created for each locus by using 300uL of Taq mix (2x) and 6uL of each primer (12 reactions total). 26uL of master mix was combined with variable amounts of DNA and water (dependent of DNA concentration from the extraction) for a final 50uL volume. In total, we had 4 corals with 2 replicates each, resulting in 8 total reactions. The PCR temperature profiles

used for the ND4/MUT and COI loci was a denaturing at 94C for 3 minutes, annealing at 54C-58C for 1.5 minutes, and extension at 72C for 1 minute for 30 cycles. For the 28s loci, the profile temperature was a denaturing at 94C for 1 minute, annealing at 58C for 1 minute, and extension at 72C for 1 minute for 35 cycles.

PCR Product Purification

Each PCR product was purified using the QIAquick PCR purification protocol. We added 5 volumes of Buffer PB (250 uL) to 1 volume of the PCR sample (50 uL) and mixed. We then placed this mixed sample into a QIAquick spin column in a provided 2 ml collection tube and centrifuged at 13600 x g for 30–60s (note, every centrifugation was done at this speed). Flow-through was discarded and the QIAquick column was placed back into the same tube. This was then washed with 750uL of Buffer PE and centrifuged for 30–60 s. This flow-through was discarded and the same column was centrifuged for an additional 1 min. We waited 5 minutes to ensure all ethanol had evaporated and afterwards the QIAquick column was then placed in a clean 1.5 ml microcentrifuge tube and eluted with either 50 µl Buffer EB or 30 uL of nuclease-free water (Table 2).

Sanger Sequencing and Phylogenetic Analysis

We utilized the economy sequencing services at the University of Alberta's Molecular Biology Services Unit. The purified PCR products were standardized to have at least 7.5 uL of 22.5 ng/uL PCR product and 2.5uL of 1uM primer (done for each sample and each individual loci). It should be noted that the DNA concentrations were varied among samples and water was used to supplement the final volume of 10uL for each sample. Retrieved sequences were organized and analyzed by the GENEious prime software (version 2023.0.4). Received nucleotides sequences were trimmed, specifically regions that had more than a 5% chance of an error per base.

These sequences were also aligned for each coral sample and its respective loci. A multiple alignment was then done on each genetic loci which included the sequences of each coral that was amplified by that respective primer. This alignment had directions automatically determined with a global alignment with free end gaps and a cost matrix of 65% similarity. Afterwards, phylogenetic trees were constructed for these multiple-aligned sequences with the Tamura-Nei genetic distance model, with bootstrap values of 100.

Results

The DNA extraction of the corals was most successful with the animal kit when compared to the plant kit. For example, the *Anthelia* DNA extracted with the animal kit provided a concentration of 48.450 ng/uL when compared to the plant kit's 13.550 ng/uL. These DNA concentrations for the animal kit extractions can be seen in table 1. A polymerase chain reaction was performed on the extracted DNA for each locus with its respective primers. These were subsequently run on a gel as shown in figure 1. It should be noted that certain temperatures, loci, and primers proved difficult in amplifying select regions in specific corals. Specifically, the soft coral *Anthelia* was not sufficiently amplified with the ND4 and MUT primers, resulting in a faint band. Another aspect to note is that the *Ricordea* soft coral was not amplified by any of our primers. A gradient PCR was performed for the problematic ND4/MUT locus for *Anthelia* and an annealing temperature of 55C was found to provide better amplification when compared to our pre-set 58C (Figure 2). However, neither the *Ricordea* samples or the ND4/MUT locus for *Anthelia* were amplified. For the PCR purification, we saw a massive increase in DNA concentration when the final product was eluted with 30uL of water instead of the normal 50uL of buffer (Table 2). The 30uL elution samples were prepared and sent for sequencing, with expected and high-quality sequences retrieved. These sequences were then trimmed and aligned, creating

consensus sequences for each coral and genetic locus, which were then all combined to form a multilocus sequence (DNA barcode). The genetic locus of 28s had more differences in base pairs between the sequences when compared to the COI locus; the multilocus barcode also has almost double the amount of base pair differences than the 28s locus (Table 3). Individual aligned sequences of each locus were also searched in the BLAST database to determine the species, with *Anthelia glauca, Sarcophtyon trocheliophorum,* and a *Sinularia spp.* being the closest matches (>95% sequence similairty). Phylogenetic trees were constructed as intended and the following relationships were analyzed as seen in figure 4.

Discussion

As expected, the animal kit provided much higher DNA concentrations due to its reagents being better suited to digesting animal tissue and proteins. However, an exception could be seen with the *Ricordea* species. This species of soft coral produced an excess of slime that was difficult to digest with this animal kit. This was reflected in the poor results obtained in the DNA extraction and the lack of PCR amplification in any of the three loci. According to the 2X CTAB DNA Isolation Procedure, using 2-mercaptoethanol could resolve this issue found in slimy marine invertebrates such as *Ricordea* (Lab & CR Renicke, 2018). The gel electrophoresis of the PCR amplification provided the expected results of ~1000 base pairs for COII8068F/COIOCTR, ~870 base pairs for ND42599F/MUT3458R, and ~775 base pairs for 28sR/28sF (Mcfadden et al., 2011). However, none of these loci were amplified in *Ricordea*, which was most likely due to the problematic DNA extraction. *Anthelia DNA* was also not properly amplified for the ND42599F/MUT3458R locus. However, the method in which the coral's DNA is preserved and extracted may hold less significance than the gene region being utilized (Pratte and Kellogg 2021). For example, the use of our animal and plant extraction kits may not hold as much influence as the

gene regions we chose (COI, ND4, etc.) The inclusion of a mock community can provide an evaluation of any bias and contamination. This result could also be attributed to the nuances of octocoral PCR. Octocorals require that we anneal the primers at the species-specific temperature (Stemmer et al., 2013). This required trial and error regarding the efficacy of certain primers and annealing temperatures. Therefore, we conducted a gradient PCR to determine an optimal temperature. Initially, the PCR was done at an annealing temperature of 58C for this locus, however, the annealing temperature range in McFadden et al. (2011) indicated 54C-58C for the ND4/MUT loci. The gradient PCR resulted in the annealing temperature of 55C to be chosen, as it provided the brightest band on the gel, indicating an improved amplification (Figure 2). Nonetheless, this new temperature was still not sufficient in providing enough DNA for the sanger sequencing. Therefore, we can assume that the locus of ND42599F/MUT3458R is problematic for our species of *Anthelia*. This could be most likely attributed to a primer mismatch in the *Anthelia* genome as this primer worked for every other coral sample. The genome where this primer was to anneal may have been altered or mutated, preventing the primer from annealing. To circumvent this, the ND4/MUT primers could be redesigned for *Anthelia* specifically, but was not done due to time constraints.

The initial rounds of PCR purification resulted in insufficient DNA concentrations. This could be attributed to the initial elution of our purified PCR products with 50uL of buffer. The QIAquick PCR purification protocol suggested to use 30uL of water for improved DNA concentrations. The second round of purification was done with this step and resulted in greatly improved DNA concentrations (Table 2). These samples were then prepared for sanger sequencing, with the retrieved sequences being of high quality. The sequences were trimmed at regions where base error rate were 5% or higher and respective sequences were then aligned. As

predicted from previous literature, the COI region was the most conserved loci among the soft corals, as it had the least number of non-identical base pairs between sequences (Table 3). This can be attributed to the unusually slow rates of mitochondrial gene evolution (~100x slower) which have caused the species-specific DNA barcodes using this COI locus for octocorals to lag behind other groups (McFadden et al., 2014). This egregiously slow rate can leave little room to differentiate between species as they can still share similar mtDNA sequences in this COI locus. The most variable region was the 28s ribosomal DNA sequence, as it had the greatest number of non-identical base pairs between sequences, various SNPs, and a gap that may indicate divergence in the 285-299 base pair region in the *Anthelia sp.* 28s locus (table 3 and figure 3). This gap in the 28s region provides discriminatory power in differentiating between species and can be attributed to an evolutionary event such as an insertion or deletion, where this sequence does not contain base pairs that are prevalent in the sequences of other species.

The species that was identified in our *Anthelia* sample's BLAST search corresponded to *Anthelia glauca* with a >95% match. The phylogenetic analysis also reveals that multiple genetic loci provide a better resolution when compared to each individual locus, as the multilocus barcode has almost double the number of base pairs that differ between the species' loci (Table 3). In general, more base pairs differing between the sequences and species results in stronger discriminatory power in identifying species. We can also see that *Nepthea* is related more to *Sarcophtyon* than it is to *Anthelia glauca* and vice versa (Figure 4). We can be confident in this analysis as the bootstrap values of 100 indicate 100% statistical confidence in the relation between the species. In terms of data, we can see the number of nucleotides that differ between each species by using this multilocus DNA barcode we have constructed (table 3). This provides evidence for the credibility and resolution of this DNA barcode as these differing nucleotides

(SNPs) are used to help determine if this barcode is sufficient in identifying a species and its relation to other species (Quattrini et al., 2019). Research has shown that species become more distinct from one another when multiple, independent loci are used in identifying them (Liu et al., 2017). More genetic loci result in greater separation, precision, and discriminatory power within or between species. By looking at the data in table 3, we can clearly this trend as we can observe more differences in nucleotides (SNPs) in the multilocus barcode when compared to each individual locus. However, we were not able to achieve a barcode of 3 genetically distinct loci for our *Anthelia* species. This is due to the ND4/MUT locus not being standardized, containing highly conserved priming sites, or having any reliable DNA amplification/sequencing for this species (Valentini et al., 2009). We assumed that we had the *Anthelia glauca* or *Anthelia elongata* species initially in sampling, based on limited morphological information as they are both morphologically similar. After the molecular analysis, the result was also *Anthelia glauca*. Therefore, we can assume that the COI and 28s loci are sufficient in providing a clear resolution on a species identification, even if the ND4/MUT locus was not being amplified in this species. Replacing the ND4/MUT locus and even adding an additional genetic locus could provide a better resolution when comparing this species to a closer relative. A barcode of three genetic loci was successfully made for the *Nepthea* species, however, the resulting molecular analysis revealed a completely different species named *Sinularia spp.* This could be due to this species of *Nepthea* not being categorized in the NCBI database, potentially indicating a new addition into the database. On the other hand, we believe that it is most likely a *Sinularia spp.* given the very high matches between the sequences in the BLAST search (>98%) and an analysis of its sclerites would be suggested to confirm this identification. For the *Sarcophyton* species, it is most likely *Sarcophyton trocheliophorum*, as the individual sequences barcodes of COI, 28s, and ND4/MUT provide a

~98% match between the species we sampled and *Sarcophyton trocheliophorum.* The *Ricordea* species was omitted from our analysis due to its faulty DNA extraction and analysis.

Coral reefs are hosts to a vast array of species that comprise a diverse and complex marine environment. Corals comprise an integral part of marine communities and play an important role in marine habitats. The only way to preserve and protect such habitats is to know which species are endangered and this can only be achieved with proper identification techniques, such as DNA barcoding. Global warming has taken a toll on coral reef diversity and health. There have been unprecedented mass coral bleaching events since the 1980s (Hughes et al., 2017). The thermal stress caused by increased global temperatures can disrupt the soft coral and its symbiotic relationship with dinoflagellates, causing loss in color (bleaching). Prolonged coral bleaching results in the death of corals and leaves behind muddy rubble which is insufficient for coral colony growth or identification. However, individual coral taxa are bleached to varying extents. This can cause a shift in coral communities, where various coral species are overtaken by fewer, more resistant coral species (such as *Anthelia*) (Webb et al., 2021). The promotion of these tolerant opportunists can eliminate other crucial species which are more sensitive to bleaching. For this reason, it is of utmost importance to identify and protect vulnerable coral species and their reefs. In our case, species such as *Anthelia* can be problematic given their capacity of recolonization, potentially making them invasive. Furthermore, when this is combined with the muddy rubble left behind after severe bleaching and the subsequent destruction of coral reefs, it can cause drastic shifts in reef diversity and health. Precious corals have also been used to produce jewelry and ornaments (Lendvay et al., 2020). Such corals demand a high price and fuel a global multi-million-dollar industry. This lucrative and profitable industry has caused many coral beds to be overharvested, severely disrupting marine habitats. These colonies of precious corals also take

lengthy times to grow and recover. This has prompted strict regulation in coral jewelry and has marked endangered species of corals as illegal to harvest. However, we have seen how difficult it is to identify a coral based only on its morphology. This difficulty in only exacerbated when the coral is polished and carved into jewelry. This requires the use of DNA barcoding for the accurate molecular identification of corals and is of paramount importance for both conservation and the enforcement of precious coral trade regulations. DNA barcoding is a cumulative process, whether it be identifying new species or protecting endangered ones. Each coral genome that is sequenced can add to the barcoding database and allows us to infer increasingly broader ranges of taxa. There are millions of new potential species out there and many that require conservation or monitoring. Future research is very inclusive when it comes to this topic. Any taxa, whose sequence is identified and submitted, can contribute to the overarching barcoding databases. The larger these databases become, the more efficient this process will be. In our study, we have contributed an understanding to the PCR and identification of *Anthelia glauca* and *Sarcophtyon trocheliophorum*. However, as stated before, we have yet to develop a universal barcode that follows all the criteria set for DNA barcodes. This itself is a daunting task that will require much trial and error. Research can also be directed into utilizing new genetic loci with their respective pairs of primers in different species of corals.

Acknowledgements

I would like to thank Dr. Joshua Miller, Dr. Ross Shaw for the amazing tutelage and guidance they provided me with this project. I also thank Kathy Davies for her support and co-ordination of this program and my honors degree. I would also like to thank Jennifer Bubenko and Victoria Bowles for their assistance regarding technique and equipment in the lab.

References

Hughes, T., Kerry, J., Álvarez-Noriega, M. *et al.* Global warming and recurrent mass bleaching of corals. *Nature* **543,** 373–377 (2017). https://doi.org/10.1038/nature21707

Jahajeeah, D., Bhoyroo, V., & Ranghoo-Sanmukhiya, M. (2021). An assessment of soft coral community (Octocorallia; Alcyonacea) around Mauritius and Rodrigues Islands—New records of soft corals. *Regional Studies in Marine Science*, *47*, 101976. https://doi.org/10.1016/j.rsma.2021.101976

Knowlton, N. (2001). The future of coral reefs. *Proceedings of the National Academy of Sciences*, *98*(10), 5419–5425. https://doi.org/10.1073/pnas.091092998

Koido, T., Imahara, Y., & Fukami, H. (2019). High species diversity of the soft coral family Xeniidae (Octocorallia, Alcyonacea) in the temperate region of Japan revealed by morphological and molecular analyses. *ZooKeys*, *862*, 1–22. https://doi.org/10.3897/zookeys.862.31979

Lab, P., & CR Renicke, C. (2018). *2×CTAB Protocol for (simultaneous) DNA Isolation from Aiptasia and Symbiodinium v1* [Preprint]. https://doi.org/10.17504/protocols.io.qx6dxre

Lendvay, B., Cartier, L. E., Gysi, M., Meyer, J. B., Krzemnicki, M. S., Kratzer, A., & Morf, N. V. (2020). DNA fingerprinting: An effective tool for taxonomic identification of precious corals in jewelry. *Scientific Reports*, *10*(1), 8287. https://doi.org/10.1038/s41598-020-64582-4

Liu, J., Jiang, J., Song, S., Tornabene, L., Chabarria, R., Naylor, G. J. P., & Li, C. (2017). Multilocus DNA barcoding – Species Identification with Multilocus Data. *Scientific Reports*, *7*, 16601. https://doi.org/10.1038/s41598-017-16920-2

McFADDEN, C. S., Benayahu, Y., Pante, E., Thoma, J. N., Nevarez, P. A., & France, S. C. (2011). Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources*, *11*(1), 19–31. https://doi.org/10.1111/j.1755-0998.2010.02875.x

McFadden, C. S., Brown, A. S., Brayton, C., Hunt, C. B., & van Ofwegen, L. P. (2014). Application of DNA barcoding in biodiversity studies of shallow-water octocorals: Molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs*, *33*(2), 275–286. https://doi.org/10.1007/s00338-013-1123-0

McFadden, C. S., Gonzalez, A., Imada, R., Shi, S. S., Hong, P., Ekins, M., & Benayahu, Y. (2019). Molecular operational taxonomic units reveal restricted geographic ranges and regional endemism in the Indo-Pacific octocoral family Xeniidae. *Journal of Biogeography*, *46*(5), 992–1006. https://doi.org/10.1111/jbi.13543

McFadden, C. S., Sánchez, J. A., & France, S. C. (2010). Molecular Phylogenetic Insights into the Evolution of Octocorallia: A Review. *Integrative and Comparative Biology*, *50*(3), 389–410.

Neigel, J., Domingo, A., & Stake, J. (2007). DNA barcoding as a tool for coral reef conservation. *Coral Reefs*, *26*(3), 487. https://doi.org/10.1007/s00338-007-0248-4

Pratte, Z. A., & Kellogg, C. A. (2021). Comparison of preservation and extraction methods on five taxonomically disparate coral microbiomes. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.684161

Quattrini, A. M., Wu, T., Soong, K., Jeng, M.-S., Benayahu, Y., & McFadden, C. S. (2019). A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evolutionary Biology*, *19*(1), 116. https://doi.org/10.1186/s12862-019-1427-y

Stemmer, K., Burghardt, I., Mayer, C., Reinicke, G. B., Wägele, H., Tollrian, R., & Leese, F. (2013). Morphological and genetic analyses of xeniid soft coral diversity (Octocorallia; Alcyonacea). *Organisms Diversity & Evolution*, *13*(2), 135–150. https://doi.org/10.1007/s13127-012-0119-x

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution*, *24*(2), 110–117. https://doi.org/10.1016/j.tree.2008.09.011

Webb, A. E., de Bakker, D. M., Soetaert, K., da Costa, T., van Heuven, S. M. A. C., van Duyl, F. C., Reichart, G.-J., & de Nooijer, L. J. (2021). Quantifying functional consequences of habitat degradation on a Caribbean coral reef. *Biogeosciences*, *18*(24), 6501–6516. https://doi.org/10.5194/bg-18-6501-2021

Figure legends

Figure 1: A 2% gel electrophoresis of candidate coral DNA that was amplified with primers for their respective loci. Two samples for each of the corals *Anthelia* (Anth), *Nepthea* (Nept), *Ricordea* (Ricor), and *Sarcophyton* (Sarc) were amplified for the COII8068F/COIOCTR (COI), ND42599F/MUT3458R (ND4/MUT), and the 28sR/28sF loci (28s). A 1kB ladder was used and a negative control without DNA was also loaded onto the gel (-). Each lane has 2uL of sample loaded. Samples were not directly stained as the PCR dye was sufficient for the gel electrophoresis.

Figure 2: A 2% gel electrophoresis of the gradient PCR performed on the *Anthelia* ND42599F/MUT3458R (ND4/MUT) locus. Each lane has Anthelia DNA that was each amplified at a different temperature. A 1kB ladder was used and a negative control without DNA was also loaded onto the gel (-). Each lane has 2uL of sample loaded. Samples were not directly stained as the PCR dye was sufficient for the gel electrophoresis.

Figure 3: Nucleotide alignment of the 28s locus for Sarcophyton (3A/3B), Anthelia (A1/A2), and Nepthea (1A/1B). Alignment and consensus were generated in the GENEious prime software. The blue highlights are altered nucleotide bases that indicate potential SNPs. The dashed/red regions indicate a potential divergence in sequences for the respective samples.

Figure 4: Phylogenetic trees constructed using the aligned sequences of each soft coral *Sarcophyton* (3A/3B), *Anthelia* (A1/A2), *Ricordea* (2A/2B) and *Nepthea* (1A/1B) for each genetic locus. These trees were constructed in GENEious using the Tamura-nei method. *Note, not enough data to construct ND4/MUT phylogenetic tree

Figures and Data

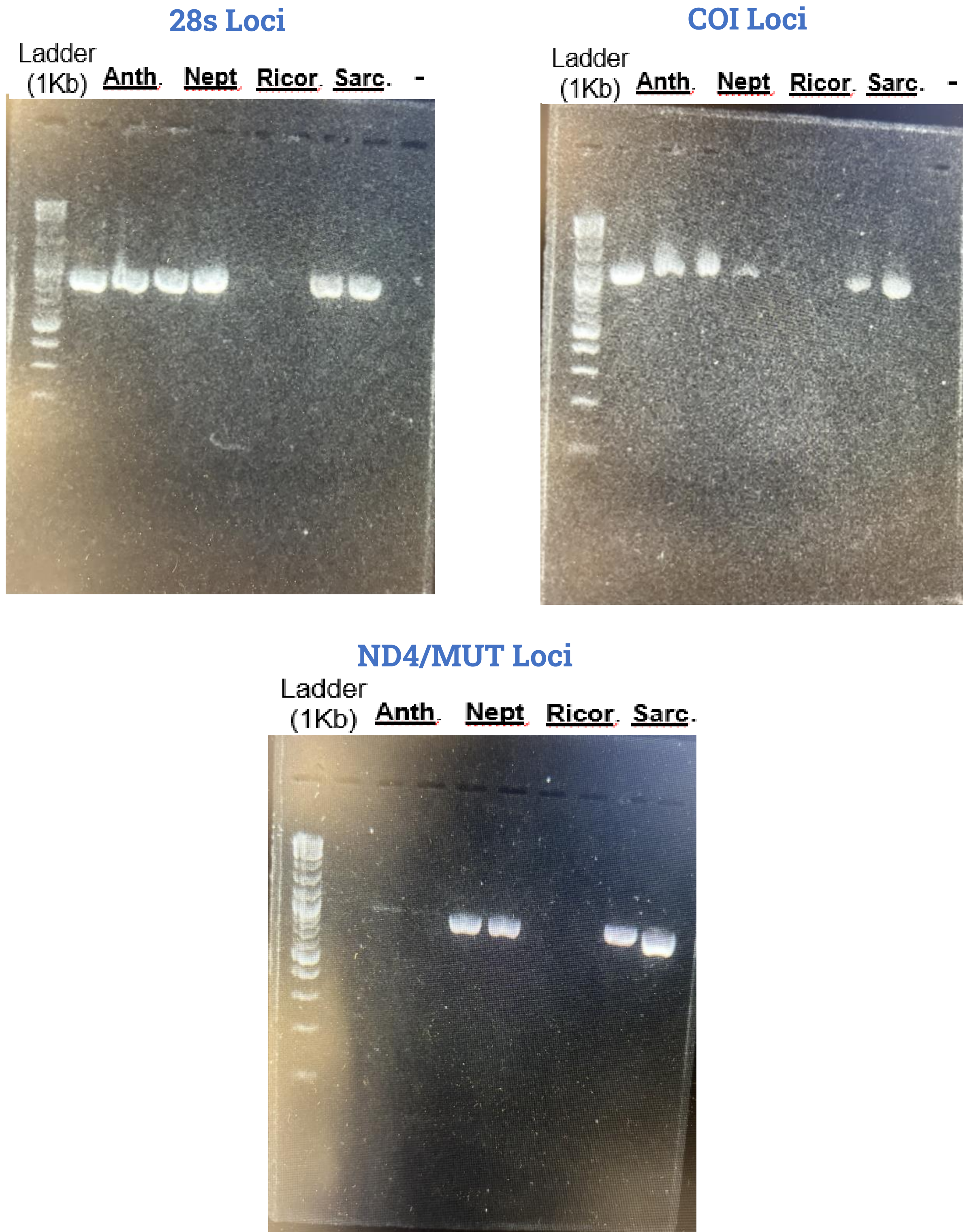


Figure 1: Gel electrophoresis of candidate coral DNA PCR for *Anthelia* (Anth), *Nepthea* (Nept), *Ricordea* (Ricor), *Sarcophyton* (Sarc) in the stated genetic loci

Gradient PCR: ND4/MUT Loci for Anthelia

- 51°C 52 53 54 55 56 57 58 Ladder (1Kb)

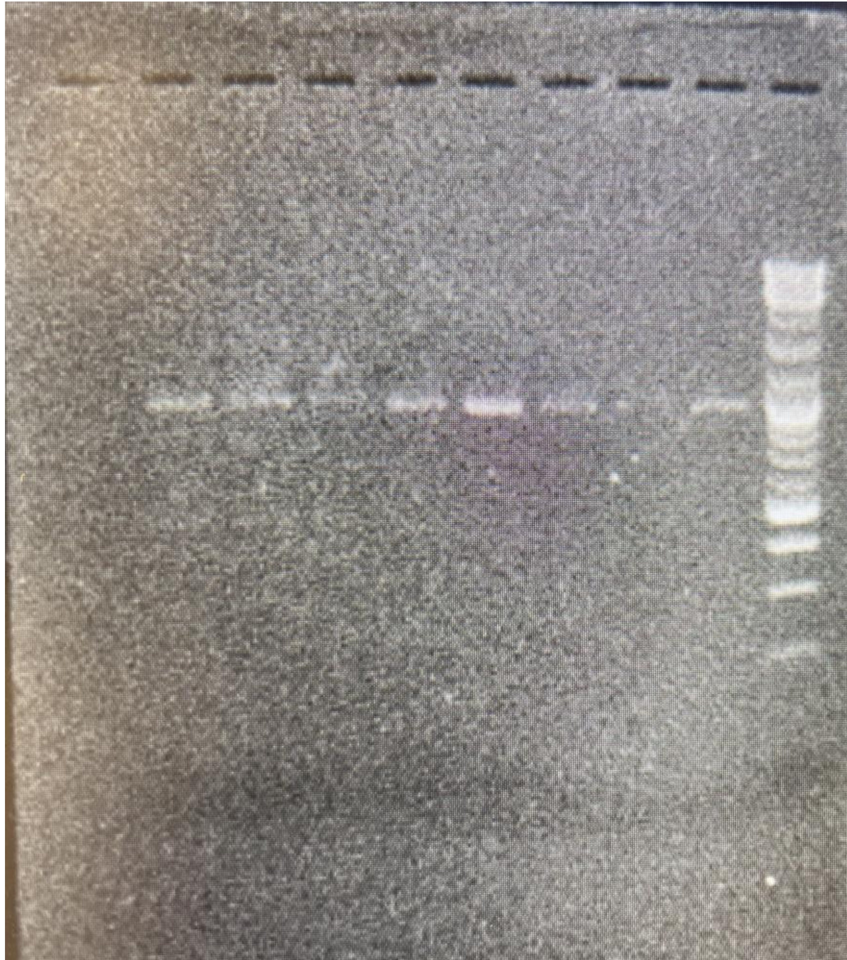


Figure 2: Gel electrophoresis of the gradient PCR for the ND42599F/MUT3458R (ND4/MUT) locus in *Anthelia*.

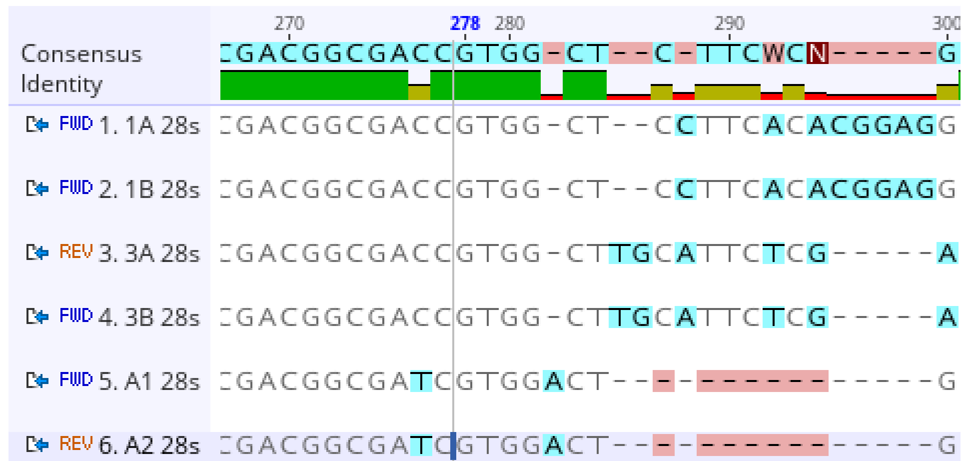
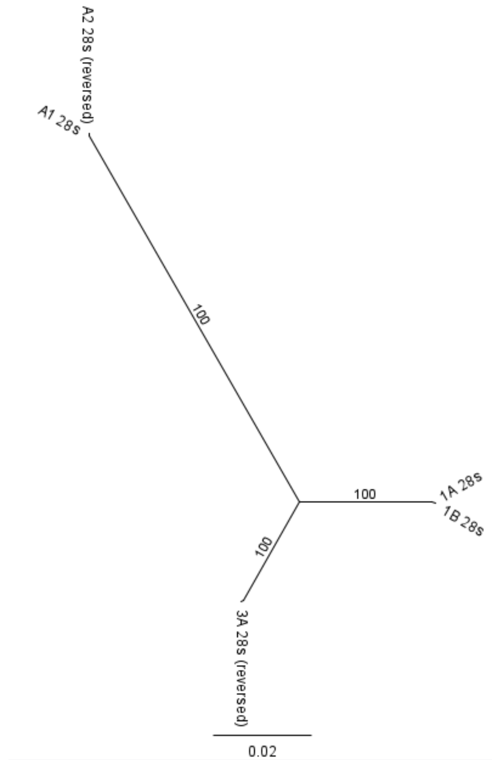
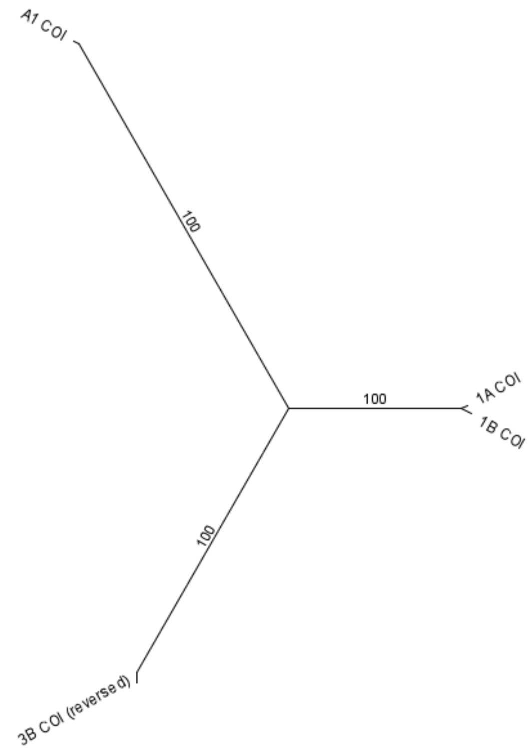


Figure 3: Nucleotide alignment of the 28s locus for *Sarcophyton* (3A/3B), *Anthelia* (A1/A2), and *Nepthea* (1A/1B)

28s Locus



COI Locus



COI + 28s +ND4/MUT

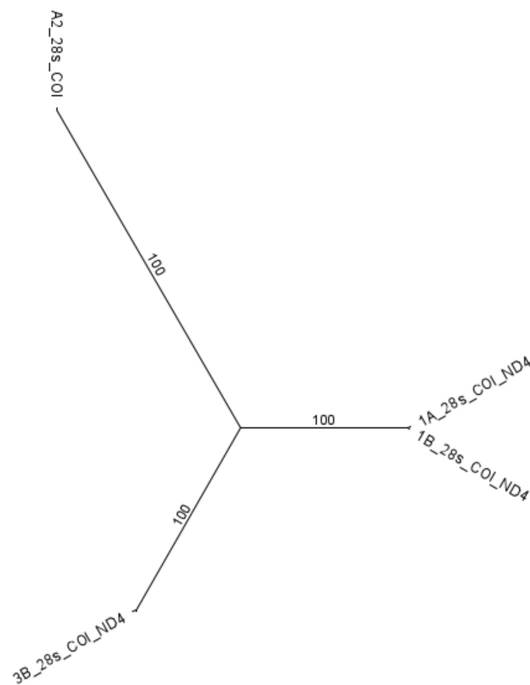


Figure 4: Phylogenetic trees constructed using the aligned sequences of each soft coral *Sarcophyton* (3A/3B), *Anthelia* (A1/A2), *Ricordea* (2A/2B) and *Nepthea* (1A/1B) for each genetic locus.

Table 1: The concentrations of DNA extractions of the candidate soft corals using an animal extraction kit

	Nepthea 1A	**Nepthea 1B**	**Ricordea 2A**	**Ricordea 2B**	**Sarcophyton 3A**	**Sarcophyton 3B**	**Anthelia 1**	**Anthelia 2**
Concentration (ng/uL)	91.35	42.2	36.2	29.9	60.75	113.3	79.7	48.45

Table 2: Concentrations following PCR purification for the *Sarcophyton* (3A/3B), *Anthelia* (A1/A2), *Ricordea* (2A/2B) and *Nepthea* (1A/1B). Round 1 corresponds to the elution done at 50uL of buffer whereas round 2 corresponds to elution done at 30uL of water. *Ricordea not included in final round due to problematic DNA extraction

	Round 1		**Round 2**	
Sample	**Concentration (ng/ul)**	**Buffer volume (uL)**	**Water volume (uL)**	**Concentration (ng/ul)**
2A COI	4.6	50	30	-
2B COI	2.85	50	30	-
3A COI	12.05	50	30	32.4
3B COI	17.1	50	30	59.45
1A 28s	20.45	50	30	118.6
1B 28s	20.2	50	30	139.35
2A 28s	3.2	50	30	-
2B 28s	7.95	50	30	-
3A 28s	22.85	50	30	64.35
3B 28s	24.65	50	30	81.5
A1 COI	16.95	50	30	85.9
A2 COI	11.5	50	30	84.2
1A COI	10.4	50	30	27.35
A2 28s	21.35	50	30	108.45
1B COI	8.45	50	30	17.6
A1 28s	24.65	50	30	145.45
1A NDMUT	12.85	50	30	46
3B NDMUT	20.75	50	30	62.4
A2 NDMUT	3.5	50	30	18.2
A1 NDMUT	3.85	50	30	19.5
3A NDMUT	18	50	30	69.35
1B NDMUT	13.9	50	30	42.25

Table 3: The number of bases that are not identical between the sequences in the respective loci of soft corals *Sarcophyton* (3A/3B), *Anthelia* (A1/A2), *Ricordea* (2A/2B) and *Nepthea* (1A/1B)

28sR/28sF (28s).

	1A 28s	1B 28s	3A 28s	3B 28s	A1 28s	A2 28s
1A 28s		1	58	58	107	106
1B 28s	1		57	57	106	106
3A 28s (reversed)	58	57		3	108	108
3B 28s	58	57	3		108	110
A1 28s	107	106	108	108		4
A2 28s (reversed)	106	106	108	110	4	

COII8068F/COIOCTR (COI)

	1A COI	1B COI	3A COI	3B COI	A1 COI	A2 COI
1A COI		1	51	64	68	64
1B COI	1		52	64	66	64
3A COI	51	52		21	82	78
3B COI (reversed)	64	64	21		93	91
A1 COI	68	66	82	93		4
A2 COI	64	64	78	91	4	

MUT3458R/ND42599F (MUT/ND4)

	3B MUT/ND4	3A MUT/ND4	1B MUT/ND4	1A MUT/ND4
3B MUT/ND4		7	70	66
3A MUT/ND4 (reversed)	7		65	67
1B MUT/ND4	70	65		4
1A MUT/ND4	66	67	4	

28s + COI + MUT/ND4 Barcode

	Nepthea 1	Nepthea 2	Sarcophyton 1	Sarcophyton 2	Anthelia 1	Anthelia 2
Nepthea 1		35	194	227	201	197
Nepthea 2	35		220	219	206	205
Sarcophyton 1	194	220		53	199	195
Sarcophyton 2	227	219	53		208	208
Anthelia 1	201	206	199	208		8
Anthelia 2	197	205	195	208	8	